The Effects of Speaker Tempo on Speech Intelligibility in Cross-dialectal Multi-talker Babble

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for graduation

with research distinction in Speech and Hearing Sciences in the undergraduate colleges of

The Ohio State University.

By

Emily R. Barker

The Ohio State University

May 2010

Project Advisors: Dr. Robert Fox and Dr. Ewa Jacewicz

ABSTRACT

The presence of background noise reduces one's ability to understand speech. Informational masking, the focus of this study, occurs when speech is masked by conversational babble. Recent studies indicate that linguistic interference contributes to informational masking but do not provide information on the sensitivity of listeners to the specific linguistic characteristics of the interfering babble or embedded speech. This study examines two questions in order to explore the nature of linguistic interference during speech-in-speech comprehension: How do dialect differences between embedded speech and multi-talker babble used as a masker affect the nature and extent of informational masking? Also, how does the tempo of embedded speech affect the interference of babble? In this study, listeners were presented with target sentences embedded in either multi-talker babble produced by speakers of a Midlands dialect of American English or by speakers of an Appalachian dialect of American English. Both 2-talker and 6-talker babble was created by each dialect group. The target sentences were mixed with the babble at different levels of signal to babble ratios (S/B ratios): +5, 0, and -5 dB. The target sentences were produced by either a slow talker or a relatively fast talker (both from North Carolina, speaking an Appalachian dialect of English). Participants, from Central Ohio, listened to 96 speech+babble combinations and typed the embedded sentence in response to a computer prompt. Error rates were determined on the basis of selected target words. Results related to both questions are presented and discussed.

TABLE OF CONTENTS

ACKNOWLDGMENTS

The research conducted in this study was supported by the Colleges of the Arts and Sciences and the College of Social and Behavioral Sciences at The Ohio State University. I would also like to thank Dr. Robert Fox and Dr. Ewa Jacewicz for their support on this project.

INDEX TO FIGURES

CHAPTER 1

Introduction and Literature Review

When individuals are listening to speech, competing background noises and other conversations are often present. The perception of speech can be affected by these surrounding sounds in the environment. The “cocktail party” effect describes the ability to focus one’s listening attention to segregate and understand speech, despite surrounding conditions (Hoen et al. 2007). Recent studies have taken particular interest in situations where speech is masked by conversational babble. These studies indicate that acoustic-phonetic and linguistic content (i.e. lexical and sub-lexical items and prosody) of the competing conversational babble contribute to the decrease in intelligibility of the target speaker (e.g., Simpson and Cooke, 2005; Van Engen and Bradlow, 2008). Last year, a study conducted by Victoria Cook in conjunction with Dr. Robert Fox and Dr. Ewa Jacewicz, in the Speech Perception and Acoustics Laboratories at the Ohio State University, examined specific linguistic factors that contribute to the decrease in intelligibility of the target speaker by varying the dialect of conversational babble masking the target speaker. The current study extends Cook's findings by providing information on how dialect and tempo differences of either the babble or the embedded speech affect speech-in-speech comprehension

1.1 Masking

The presence of background noise has a detrimental effects on a listener’s ability to perceive speech. Findings from studies (Rogers et al., 2004) suggest that when speech signals are presented in noise, as opposed to in quiet, listeners process speech signals differently due to energetic masking. Energetic masking influences peripheral auditory processing and occurs when competing sounds overlap in time and frequency in such a way that one or more of the sounds becomes inaudible (Brungart et al., 2001; Hoen et al., 2007).

In addition to the energetic overlapping of noise, other factors contribute to the overall masking effect (Van Engen and Bradlow, 2007). Previous studies found that a masker containing meaningful speech caused significantly more masking than expected (e.g., Carhart et al. 1969; Kidd et al. 2007). Carhart et al. (1969) described this as "perceptual" masking because it was thought to be caused by higher levels of processing (i.e. linguistic processing). This type of masking cannot be attributed solely to the physical interaction between the signal and masker, and is referred to as informational masking (Hoen et al. 2007; Brungart et al. 2001). Informational masking affects both peripheral auditory processing and central perceptual processing. When speech is embedded in conversational babble, which is the case in the present study, energetic masking occurs, but the masking effect is mainly due to informational masking (Hoen et al., 2007)

Several variables affect the amount of masking present during speech-in-speech perception, such as signal-to-noise ratio (SNR), number of talkers in the babble, and characteristics of the speakers in both the target sentences and background speech (Brungart et al. 2001). Investigating speech-in-noise perception, whether noise was speech or non-speech, found that the performance of listeners decreases as the level of noise increases (Van Engen and Bradlow, 2007). In other words, as signal-to-noise ratio decreases, listeners' performance also decreases.

The number of talkers in conversational babble also contributes to the amount of masking during speech-in-speech comprehension. As the number of speakers increases, performance among listeners decreased (Simpson and Cooke 2005). However, intelligibility of the target signal remained about the same for babble containing 6 to 128 speakers. Simpson and Cooke (2005) compared this finding to speech intelligibility in babble-modulated noise. Natural babble masked more effectively when at least 3 talkers were present. These results imply that when more talkers are in the babble, more energetic masking is present because there are fewer temporal gaps in the masker (Van Engen and Bradlow, 2007).

When examining voice characteristics, studies generally show that when background speech and target speech are less similar, intelligibility of target speech increases. For example, Brungart et al. (2001) found that performance among listeners increased when the masker and the target were of a different sex as opposed to being presented with a same-sex masker. Van Engen and Bradlow (2007) recently found that the language of babble also affects listeners' performance. The study compared intelligibility of English target sentences masked by English with intelligibility of English target sentences masking by Mandarin Chinese when presented to native English listeners. The results in this study found that linguistic factors have a greater affect on the overall masking when there are fewer speakers present and when the level of the target is equal or less than the level of the masker. The linguistic characteristics of Mandarin Chinese and English differ greatly, providing little information regarding the source of the language effect. Other variations of the babble masker, such as dialect, could provide further insight.

1.2 Dialect Perception

Acoustic-phonetic information plays an important role in speech perception across dialects (Clopper and Bradlow 2008, Labov 2006). Clopper and Bradlow (2008) examined how dialect variation affects speech-in-noise perception. In more difficult listening conditions , participants cannot effectively adapt to dialect variation in the acoustic signal and cross-dialect differences in intelligibility emerge for all listeners, regardless of their dialect.. Listeners could better comprehend speech in the General American English category with less marked dialects (e.g. New England, the West, and the Midland region). In listening conditions with a higher signal-to-noise ratio, listeners made more transcription errors in more marked dialects, such as the Mid-Atlantic, Northern, and Southern regions.

This study demonstrates the importance of acoustic-phonetic cues of dialects on the perception of speech. Cook's (2009) study expanded upon the recent findings indicating the contribution of linguistic interference to informational masking as well as the importance of acoustic-phonetic cues in speech processing.

Cook's (2009) study compared intelligibility of target sentences masked by two different dialects that vary greatly in acoustic-phonetic characteristics but do not vary systematically in terms of phonological structure or lexical items. Two sets of multi-talker babble (used as informational maskers) were created using either speakers of a Midlands dialect of American English (spoken in Central Ohio) or speakers of the Appalachian dialect of American English (spoken in western North Carolina). Target sentences that were embedded in the babble (which listeners had to identify) were produced by a central Ohio speaker. Cook (2009) found that there was a significant effect of the dialect of babble on listener performance.

1.3 Speech Rate

The tempo at which speech is produced varies among individuals. Differences in speaking tempo affect acoustic-phonetic cues of speech signals during speech perception. Many phonemic contrasts are constituted by temporal characteristics (short vs. long vowels or consonants, affricates vs. fricatives, stops vs. glides, etc.) and are certainly altered by changes in speech rate. Several studies on speech rate indicate that as speaking rate increases, intelligibility decreases (Liu and Zeng, 2006; Jones et al.,, 2007).

Although studies on the overall affect of speech tempo on perception exist, little information is available on how speech tempo influences intelligibility during speech-in-speech comprehension. The current study aims to provide more information on the relationship between speech rate and perception by examining how varying tempos of target sentences embedded in conversational babble affect listeners' ability to identify the signal.

1.4 The Present Study

The present study extends upon Cook's findings (2009). The first part of the present study uses target sentences produced by a NC speaker embedded in the same multi-talker babble from Cook's study. However, some NC talkers have significantly slower speech tempos than do Ohio talkers. It might be the case that dialect differences and speech rate differences can have separate effects (slower speech rate might actually improve listener identifications). Thus a second set of target sentences, produced by a NC talker with a relatively slow rate of speech (a rate more common for these NC talkers), were used in the second set of perception tests.

The present study will provide further insight about the source of linguistic interference in speech-in-speech comprehension. Most previous studies on linguistic interference and informational masking (e.g., Van Engen & Bradlow, 2007) only vary the language of the babble. This study provides information on how dialect and tempo of target speakers affects the interference of babble during informational masking.

CHAPTER 2

Methodology

2.1 Babble Speakers

The speech used in the babble was recorded for a previous study in the Speech Perception and Acoustic Labs and consisted of spontaneous, natural speech. For the current study, the speakers chosen for the creation of the babble included twelve male speakers between 50 and 70 years of age. Six of the twelve were speakers of the Midlands dialect of American English (spoken in Central Ohio). The other six were speakers of the Appalachian dialect of American English (spoken in western North Carolina). To avoid any influence of voice quality differences on the intelligibility of the target speech (Brungart et al., 2001), all of the speakers used in the babble creation were of the same gender and had a similar fundamental frequency.

2.2 Babble Creation

Short individual phrases of the recorded conversations were extracted from each speaker's discourse. Each phrase contained approximately 5 to 10 syllables (e.g. "quite a few more restaurants" and "not as well educated and"). These phrases were amplitude normalized and reordered using Adobe Audition 1.0 waveform editing program. The speech was made semantically anomalous in order to eliminate syntactic and semantic effects so that listener's could hear individual words but could not follow any "story" or predict words in advance by listening to the babble. The reordered recordings were then copied and added to the end to create one long sound file for each speaker, which was also amplitude normalized using Adobe Audition 1.0.

Two speakers from each dialect were chosen for creation of the 2-talker babble. Selection of these speakers was based on similarity between the voice quality of the speaker and target speaker. This was done in order to eliminate extraneous factors that may interfere with the perception experiment. With the control of a MATLAB program, these speakers were combined into a single file to create the 2-talker babble in each dialect. To create the 6-talker babble, all 6 speakers were combined. Each combined recording was then divided into 120 samples measuring 4 seconds long. The beginning of each sample included the end of the previous sample. Due to this, the repeated sections of the recording delivered different samples because the starting point of the sample was in a different location. This created 120 samples in each of the babble conditions (2-talker Ohio dialect babble, 2-talker North Carolina dialect babble, 6-talker Ohio dialect babble, and 6-talker North Carolina dialect babble).

2.3 Target Sentences

Two male speakers of the Appalachian dialect were recorded for the target speech; a faster speaker for the first set of perception tests, and a slower speaker for the second set of perception tests. The speech-rate of the fast speaker was 4.17 syllables/second and the speech-rate of the slow speaker was 2.94 syllables/second. Both speakers were born and raised in the western North Carolina area and met the same age and fundamental frequency criteria as the speakers used in the babble. The speakers read sentences from the Revised Bamford-Kowal-Bench Standard Sentence Test lists 3, 4, 6, 7, 9, 13, 18. Each list contains 16 short, simple sentences (e.g., *The book tells a story.*). All sentences are listed in Appendix A. Each sentence has 2 to 3 keywords, combined for a total of 50 keywords per list. Six of the lists were used for the experiment for a total of 300 keywords. The recordings took place in a sound-attenuating booth using a head-mounted Shure SM10A microphone positioned 1-inch from the speaker's lips. All sentences were recorded directly to a hard disk drive at a quantization rate of 16 bits and a sampling rate of 44.1 kHz and. The recordings were peak amplitude normalized before the overall amplitude was adjusted when being combined with the babble speech.

2.4 Listeners

The group of participants for this study primarily consisted of young adults who were born, raised, and lived in Central Ohio (determined to be within approximately 60 miles of Columbus, OH). Each participant was paid $10.00 for participating in the 30 to 60 minute long listening test. Participants included 32 listeners, 13 male and 19 female, ranging between the ages of 20-27. All participants reported having normal hearing. 16 of the listeners participated in the first set of perception tests that used target sentences spoken by a fast North Carolina (NC1) speaker. Of the 16 listeners exposed to the NC1 speaker, 8 completed perception tests in the 2-talker babble condition and the other 8 completed perception tests in the 6-talker babble condition. The remaining 16 listeners participated in a separate set of perception tests that used target sentences spoken by a slow North Carolina (NC2) speaker. 8 of the 16 listeners exposed to the NC2 speaker completed perception tests in the 2-talker condition, and the other 8 completed the tests in the 6-talker condition.

2.5 Procedure

For this study, two sets of perception tests were conducted in the Speech Perception and Acoustic Lab at Ohio State. The tests were administered under the control of a MATLAB program. The first set of perception tests used target sentences spoken by a fast speaker of the Appalachian dialect of American English and the second set of perception tests used target sentences produced by a slower speaker of the Appalachian dialect of American English. Both sets of perception tests followed the same experimental design (see Table 2.1). Each listener was seated in a sound attenuating booth and listened to target sentences mixed with either the 2-talker or 6-talker babble, delivered over Sennheizer HD 600 circumaural headphones at a comfortable listening level (~70 dB HL). For each testing block, the experimenter selected the number of speakers, dialect, and signal-to-babble ratio.

Each target sentence was positioned within the center 75% of each 4 second long babble segment. To prevent listeners from predicting the exact timing of the target sentences' onset, the actual position of the sentence was randomly varied within this location. The onset and offset of the mixed token was ramped from/to zero using a Hanning window over the first and last 5 milliseconds. The levels of the babble and sentence were appropriately adjusted for each trial. Participants faced a computer monitor and were asked to type the target sentence they heard into a box that appeared on the screen after each target and babble combination. Participants either completed the perception experiment with target sentences masked by 2-talker or 6-talker babble. Each participant was presented with 100 sentences total (4 sentences in the practice block followed by 6 experimental blocks containing 16 sentences each). In the first of the 6 experimental blocks, participants listened to the target sentences masked by North Carolina babble at a signal-to-babble ratio of +5 dB. For the second block, the babble was Ohio and still at a signal-to-babble ratio of +5 dB. In the third block, the listeners were presented with target sentences masked by North Carolina babble at a signal-to-babble ratio of 0 dB. The fourth block was also at a signal-to-babble ratio of 0 dB, but the target sentences were masked by Ohio babble. The fifth and sixth blocks were presented in the same order at a signal-to-babble ratio of -5 dB. The practice block featured babble that was mixed with both Ohio speech and North Carolina speech and was presented at a babble-to-signal ratio of +5 dB.

Table 2.1. Experimental design: conditions, blocks.

Condition	**Number of Talkers in Babble**	**Block 1 North Carolina Dialect**	**Block 2 Ohio Dialect**	**Block 3 North Carolina Dialect**	**Block 4 Ohio Dialect**	**Block 5 North Carolina Dialect**	**Block 6 Ohio Dialect**
1	2	S/B Ratio: +5	S/B Ratio: +5	S/B Ratio: 0	S/B Ratio: 0	S/B Ratio: -5	S/B Ratio: -5
2	6	S/B Ratio: +5	S/B Ratio: +5	S/B Ratio: 0	S/B Ratio: 0	S/B Ratio: -5	S/B Ratio: -5

It was predicted that blocks would become more difficult as the S/B ratio decreased due to the greater level of noise and energetic masking. It was also predicted that listeners would perform more poorly when the target speech was embedded in the North Carolina babble rather than the Ohio babble due to linguistic interference. Each participant was first presented with North Carolina babble as the masker, followed by Ohio babble. Also, participants were first presented with the easiest S/B ratio (i.e. +5 dB) which was then decreased after the listener was presented with two blocks (one for each dialect of babble) at this level. All participants were presented with the blocks in the same order as described in Table 2.1.

2.6 Data Analysis

Scoring was based on the listener's ability to identify keywords provided by the Revised Bamford-Kowal-Bench Standard Sentence Test. Each sentence had 2 or 3 keywords for a total of 50 keywords per set of 16 sentences, 300 keywords in total. The scoring criteria were adapted from the Van Engen and Bradlow (2007) study investigating speech perception in babble. Participants received one point for each correctly identified keyword. Simple spelling errors and homophones were also counted correct. A word was considered incorrect if it had any added, deleted, or incorrect morphemes.

The results for each subject in each of six experimental conditions were tabulated and transformed into a percentage correct score. However, the variance of percentage (and proportion) correct scores is not independent of the mean of these scores due to ceiling and basement effects. As a result, they violate the homogeneity of variance assumption required in many statistical tests. Thus, the percentage correct scores were arcsine transformed into rationalized arcsine units or RAUs prior to data analysis (see Studebaker, 1985). On this scale, -23 RAU corresponds to 0% correct and +123 RAU corresponds to 100% correct. A modified method of scoring was also used to give subjects partial credit for answers. However, since the pattern of results was the same for both methods, only the results based on the Van Engen and Bradlow methodology are presented.

CHAPTER 3

Results

3.1 Results

The results of each subject in the six experimental conditions were tabulated and raw scores were converted into percentage correct scores.   Mean percentage correct scores are presented in Figures 3.1-3.4. However, noting back to the Methodology section, percentage correct scores violate the homogeneity of variance assumption often required for statistical analysis. Therefore, prior to the analysis of the data collected for this study, percentage correct scores were arcsine-transformed into rationalized arcsine units or RAUs (see Studebaker, 1985).

The arcsine-transformed data sets were analyzed using a 3-way repeated measures analysis of variance (ANOVA) with the within-subject factors consisting of *S/B ratio* (+5 dB, 0 dB, -5 dB) and *dialect of babble* (Ohio babble or North Carolina babble) and the between-subject factors consisting of *babble group* (2-speaker babble or 6-speaker babble) and tempo of target sentence (NC fast speaker (NC1) or NC slow speaker (NC2)). Although the RAUs were used for statistical analysis, the percentage correct scores were used in Figures 3.1-3.4

3.2 S/B Ratio

As expected, there was a significant effect of S/B ratio ($F(2,56)=478.2$, $p<.001$). Subject performance decreased as the level of the signal (the embedded target sentences) relative to the level of the babble decreased. Listeners performed extremely well in the +5 dB condition and very badly in the -5 dB condition. The mean percentages correct for the +5 dB, 0 dB, and -5 dB ratios were 93.9%, 74.3% and 39.9%, respectively. Figures 3.1-3.2 illustrate this effect. Partial eta squared (or effect size), which represents the proportion of total variability attributable to a specific factor, is .945. This means that S/B ratio account for 94.5% of the overall effect.

3.3 Babble Group

There was also a significant effect of babble group ($F(1, 28)=14.79$, $p<.001$). Listeners performed slightly better in the 2-talker babble condition (73.8%) than the 6-talker babble condition (65.9%). The comparison between Figures 3.3 and 3.4 illustrate this effect. Partial eta squared for babble group was .346, meaning 34.6% of the overall variability can be attributed to this factor. Moreover, the combined effect of babble group and S/B ratio also produced significant effects ($F(2,56)=12.36$, $p<.001$). In the +5 dB condition, 2-talker and 6-talker performances were very similar (93.7% and 94.1%). In the 0 dB condition, 2-talker and 6-talker performances differed more (78.4% and 70.2%). In the -5 dB condition, the differences in performance in the 2-talker and 6-talker conditions were even greater (49.1% and 30.6%). This is illustrated in Figures 3.1a and 3.1b, separately for the NC1 and NC2 condition. This shows that as the level of the target sentence decreases, the effect size of babble group increases. The partial eta squared for babble group and S/B ratio was .306 meaning 30.6% of the total variability can be attributed to the combination of babble group and S/B ratio.

3.4 Babble Dialect

One main interest in this study is the effect of babble dialect upon listener performance. There was a significant effect of the dialect of babble upon listener performance ($F(1,28)=13.17$, $p<.001$). The percentage of correct responses was lower when the sentences were embedded in Ohio babble (67.9%) than in North Carolina babble (70.8%). Parial eta squared for babble dialect was .320. This means that babble dialect accounted for 32.0% of the overall effect.. Additionally, there was a significant dialect by target sentence tempo interaction ($F(1,28)=14.2$, $p<.001$). This significant interaction results from the fact that dialect had a much greater effect on performance in the NC2 condition than the NC1 condition. Mean percentages correct in the slow speaker condition for Ohio babble and North Carolina babble were 67.9% and 74.7% respectively, while the mean percentages correct in the fast speaker condition for Ohio babble and North Carolina babble were much closer (68.0% and 66.9% respectively). Partial eta squared for the interaction between babble dialect and target sentence tempo was .337, meaning the interaction of these two factors account for 33.7% of the total variability. The influence of babble dialect on listener performance is illustrated in Figure 3.2 - 3.4.

3.5 Target Sentence Tempo

There was no significant effect of tempo of target sentences ($F(1, 28)=3.3$, $p>.05$) although listeners showed slightly better performance in the NC slow speaker (NC2) condition (71.3%) than in NC fast speaker (NC1) condition (65.0%). However, there was a significant effect of target sentence tempo when combined with the effect of S/B ratio ($F(2,56)=12.37$, $p<.001$). In the fast speaker condition, the mean percentages correct for +5 dB, 0 dB, and -5 dB were 93.3%, 76.5%, and 32.6%, respectively. In the slow speaker condition, the mean percentages correct for +5 dB, 0 dB, and -5 dB were 94.6%, 72.1%, and 47.2%. In the +5 dB and 0 dB conditions, the performance of listeners was similar regardless of the tempo of the target sentence. In the -5 dB condition, on the other hand, listeners performed nearly 15% better in the slow speaker condition than the fast speaker condition. Partial eta squared for the combined factors of target sentence tempo and S/B ratio was .306, meaning 30.6% of the overall effect can be attributed to the interaction of these two factors. The effects of target sentence tempo are illustrated by comparing Figures 3.1a, 3.2a, 3.3a, and 3.4a to Figures 3.1b, 3.2b, 3.3b, and 3.4b.

In addition, the interaction between dialect of the babble, S/B ratio, and tempo of the target sentence also produced significant effects ($F(2, 56)=5.03$, $p=.010$). Partial eta squared for the combination of these three effects was .152, meaning around 15.2% of the total variability is attributable to the interaction of babble dialect, S/B ratio, and tempo of the target sentence. This is illustrated in Figures 3.2a and 3.2b. None of the other interaction effects were significant.

Figure 3.1a S/B Ratio by Group – NC1

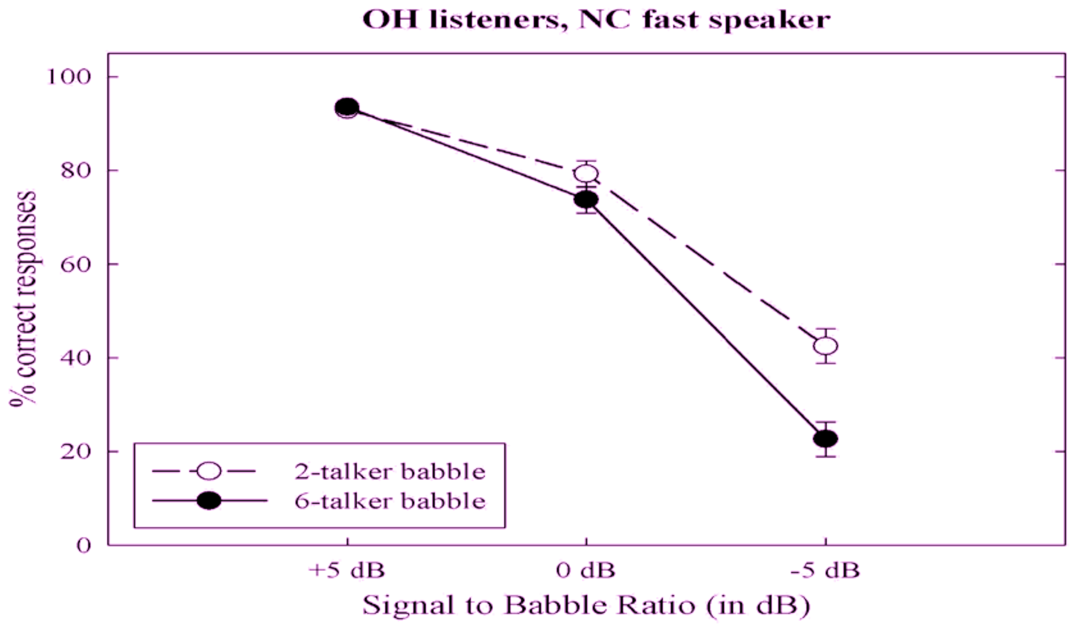


Figure 3.1b S/B Ratio by Group – NC2

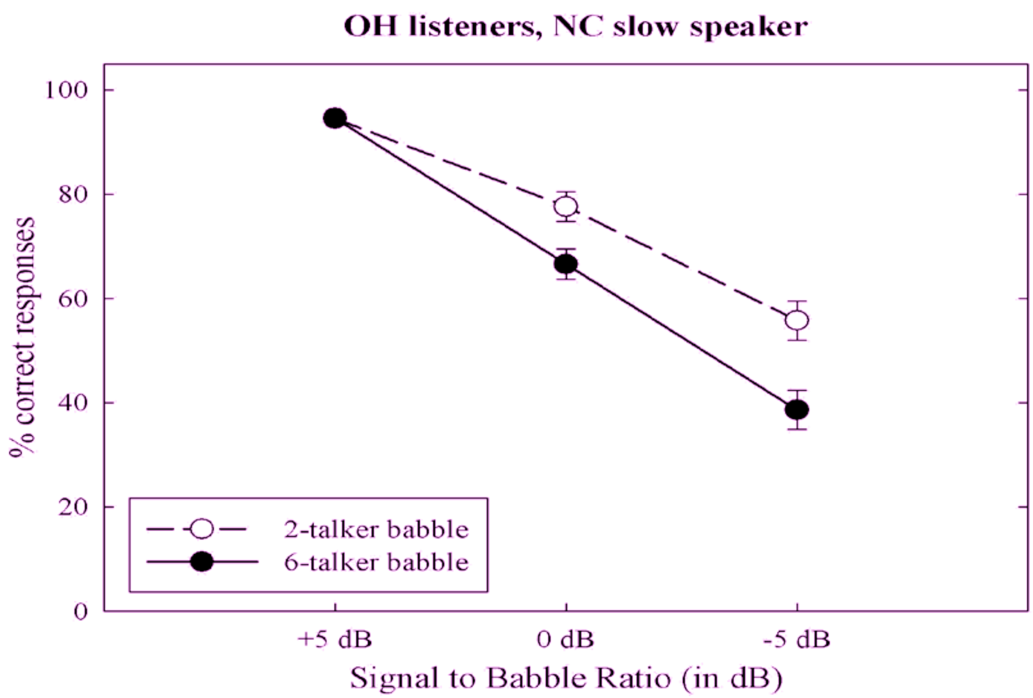


Figure 3.2a S/B Ratio by Babble Dialect – NC1

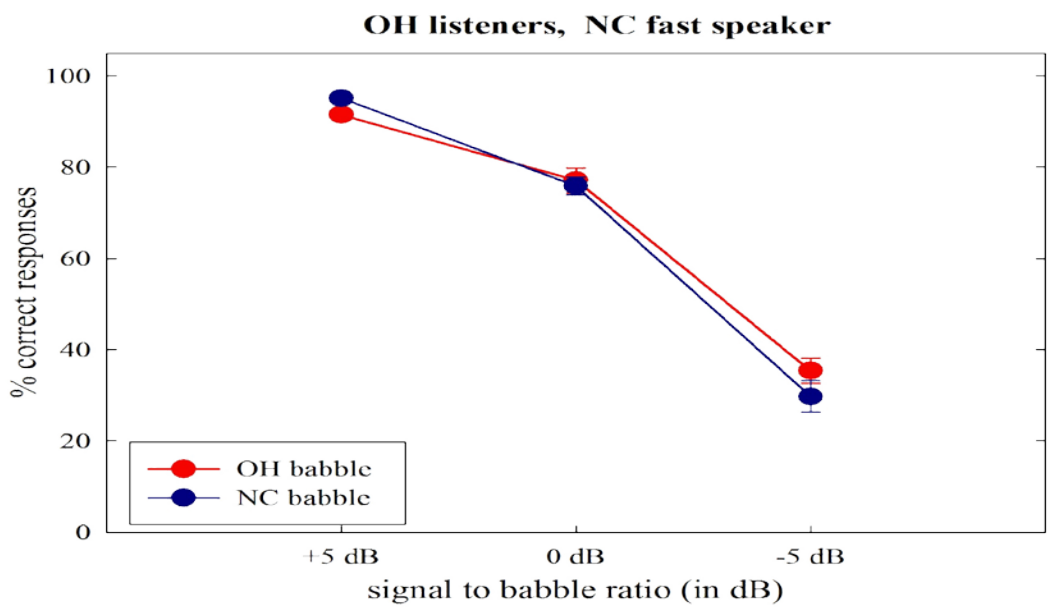


Figure 3.2b S/B Ratio by Babble Dialect – NC 2

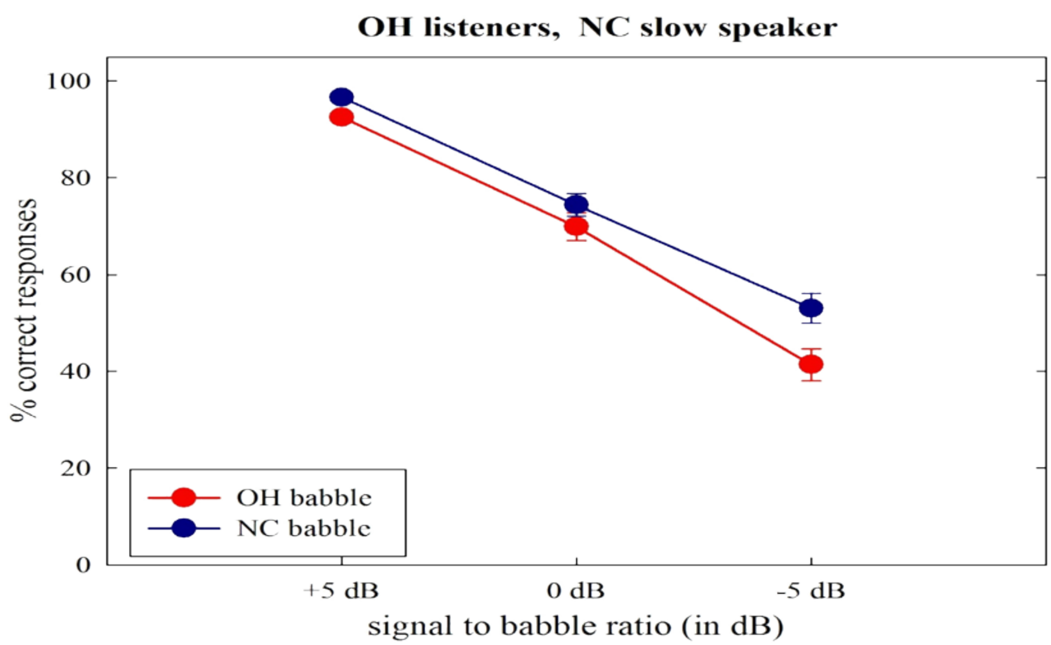


Figure 3.3a 2-Talker Babble by S/B ratio – NC1

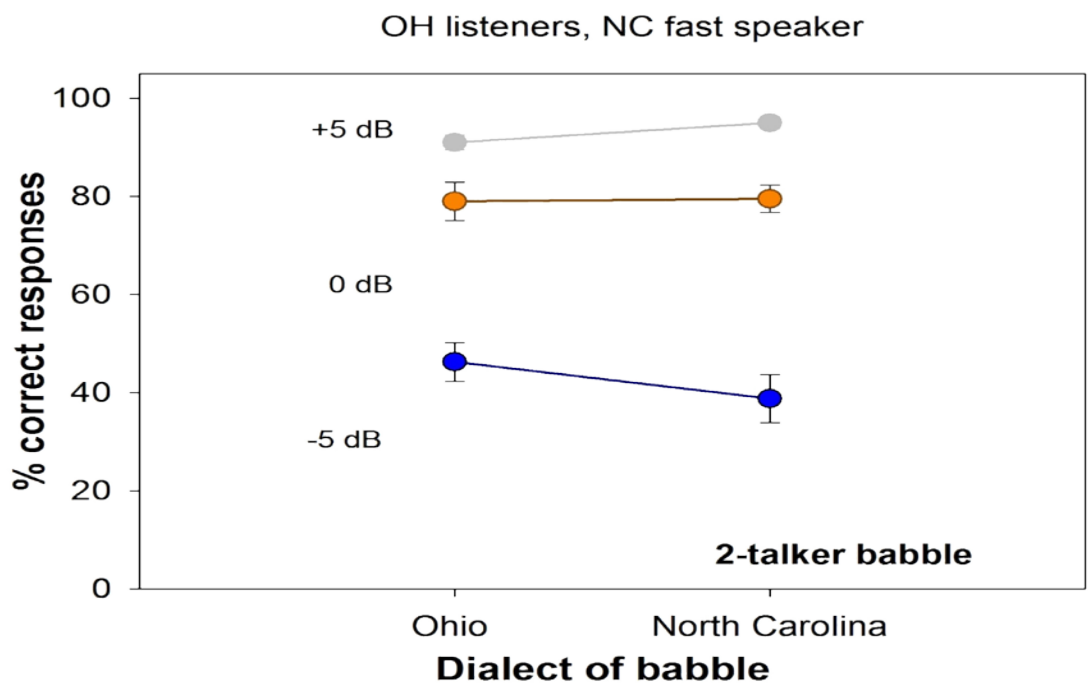


Figure 3.3b 2-Talker Babble by S/B Ratio – NC 2

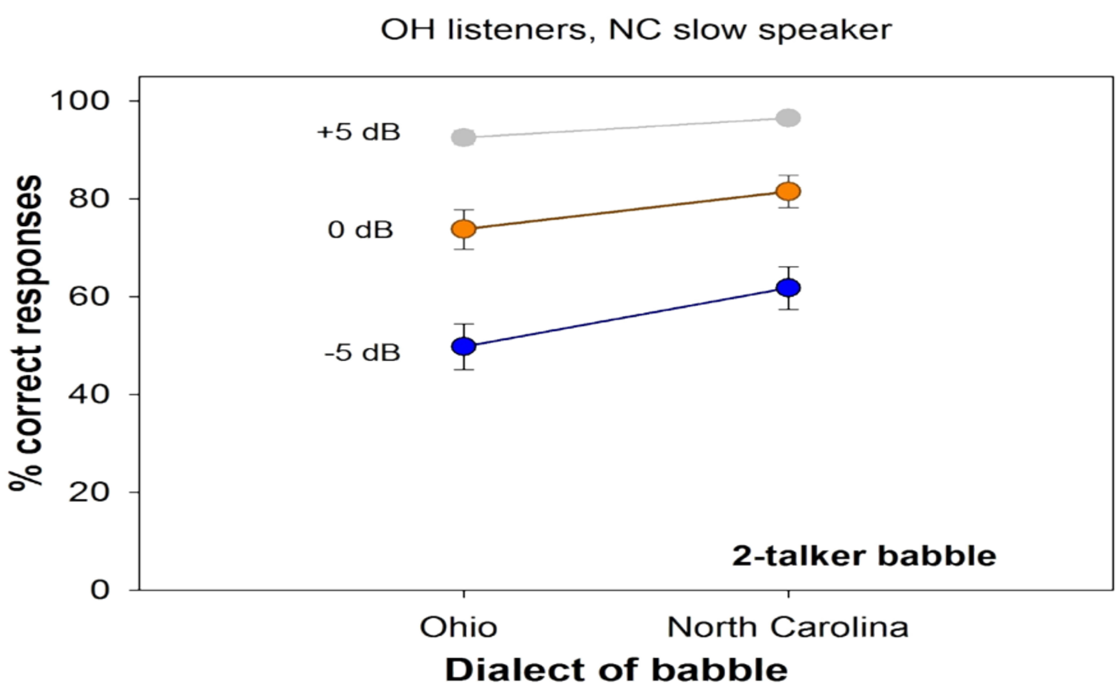


Figure 3.4a 6-Talker Babble by S/B Ratio – NC 1

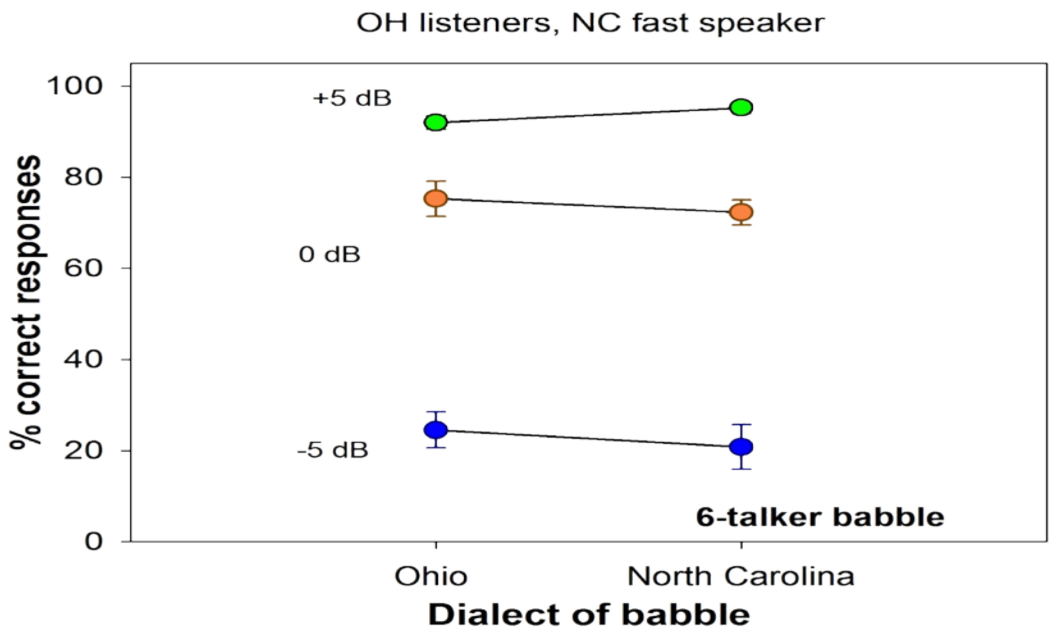
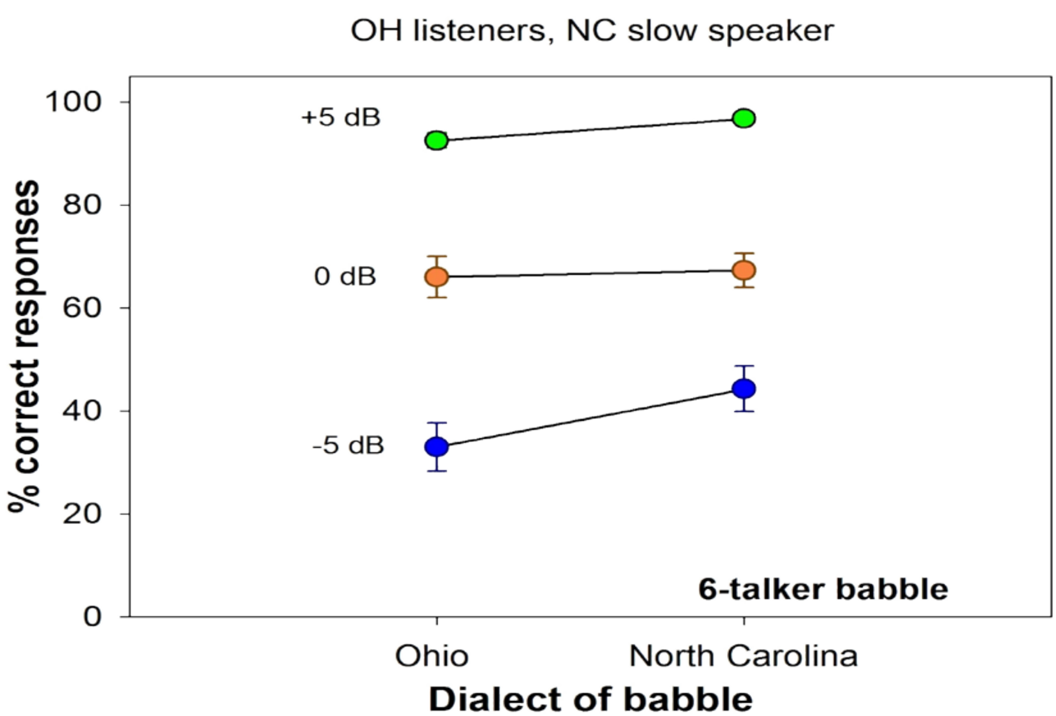


Figure 3.4b 6-Talker Babble by S/B Ratio – NC2



CHAPTER 4

Discussion and Conclusions

The ability to understand speech is reduced when it is presented in the context of conversational babble (the speech of multiple talkers added together)—this type of noise influences both peripheral auditory processing and central perceptual processing partially due to “informational masking.” Recent studies indicate that linguistic interference contributes to informational masking but do not provide extensive information on the sensitivity of listeners to the specific linguistic characteristics of the interfering babble . The present study further examines linguistic interference by comparing the effects of babble created from recordings of speakers of two distinct American English dialects on the identification of a set of target sentences. This study shows how dialect and tempo differences of the babble and the embedded speech affect speech-in-speech comprehension.

As expected, based on the results of previous studies (Van Engen and Bradlow 2007 Brungart et al. 2001), listeners' ability to perceive speech in multi-talker babble decreased as the signal-to-babble ratio decreased. In addition to S/B ratio, listener performance was also affected by babble group. Listeners performed significantly better in the 2-talker condition than the 6-talker condition. Several previous studies have also found that performance decreases as the number of talkers in the masker increases (Simpson and Cooke 2005). One specific study conducted by Hoen et. a (2007) compared word comprehension in 4-talker babble with 6-talker babble and found that comprehension rates were higher in the 4-talker condition. In the present study, there was also a combined effect of S/B ratio and babble group showing that as the level of the target sentence decreases, the number of talkers in the babble has a more significant effect on listener performance.

Of particular interest in this study is how variances in babble dialect and target sentence tempo change listeners' speech-in-speech comprehension abilities. The primary study conducted by Cook

(2009) found that Ohio babble (Midlands dialect) produced more informational masking than North Carolina babble (Appalachian dialect). Cook tested Ohio listeners exposed to Ohio target speech. Many studies have indicated that greater similarity between talkers creates greater masking of the target (Brungart et al., 2001). Studies on dialect perception have shown that listeners are sensitive to acoustic-phonetic characteristics in marked dialects. Therefore, it is possible that this effect could be due to the level of similarity between talkers, or to interference due to linguistic processing of the dialects in the background. The results of the present study provide further insight into Cook's findings. The current study also found that Ohio babble produced more informational masking than the North Carolina babble. However, in contrast with Cook's study, Ohio listeners were identifying North Carolina target speech rather than Ohio target speech. This indicates that the effect of variation in the dialect of the babble is at least partially due from listeners "tuning in" to his or her native dialect more than the non-native dialect, and not simply because of talker similarity between target speech and babble.

The tempo of the target sentence alone did not produce significant effects on speech-intelligibility although performance in the slow speaker (NC2) condition was slightly higher than in the fast speaker (NC1) condition. There was, however, a significant interaction between target sentence tempo and dialect. Dialect had a much greater effect on performance in the slow speaker condition as opposed to the fast speaker condition. Moreover, the interaction between target sentence tempo and S/B ratio was also significant. In the lowest S/B ratio (-5 dB) condition, listener's performed nearly 15% better in the slow speaker condition than the fast speaker condition. A previous study also found that listener comprehension was lower at the faster speech rate. This could be because the slower pace of the speech allows more time for the listener to pick up and process the acoustic-phonetic cues. Although it is obvious that there is an intelligibility advantage when the speaker is slow, further research should be done on how fast the speaker should be in order to still be intelligible. Moreover, future studies could also look into how dialect affects perception of speech among various tempos. For example, would a fast speaker be very difficult to understand regardless of dialect? One might predict that there is indeed a rate at which speech is no longer intelligible and this may vary across dialects. More enunciated dialects may have higher intelligibility in fast speaker conditions than others. Future studies could also investigate whether tempo variations of babble have an effect on the amount of informational masking. For example, slower speakers used in babble could result in more informational masking since the listener may be able to actually comprehend the sentences used in babble production. This could provide more information on the source linguistic interference in speech.

Conclusions

The results indicate that the dialect of the babble, the signal-to-babble ratio, and number of speakers in the babble influence listeners' ability to comprehend sentences produced in babble. The tempo of the target sentence also had a strong effect upon listener performance when combined with other attributing factors s of dialect and/or S/B ratio. Consistent with previous speech-in-babble studies, performance decreased as the signal-to-babble ratio decreased. The Ohio babble produced more informational masking than the North Carolina babble. The results of the current study support predictions based on previous studies indicating that linguistic interference contributes to informational masking (e.g., Engen & Bradlow, 2007). Unlike previous studies which varied the language of the babble, the dialects used in this study share most lexical and phonological characteristics and vary greatly only in acoustic phonetic characteristics. These findings provide further insight into the source of linguistic interference in speech-in-speech comprehension.

REFERENCES

Bench J, Kowal A, & Bamford J. (1979). The BKB (bamford-kowal-bench) sentence lists for partially-hearting children. *British Journal of Audiology*, 13(3) 108-12.

Brungart, D.S., Simpson, B.D., Ericson, M.A., and Scott, K. R. (2001). “Informational and energetic masking effects in the perception of multiple simultaneous talkers,” *The Journal of the Acoustical Society of America* 110, 2527-2538.

Carhart R, Tillman TW, & Greetis ES. (1969). Perceptual masking in multiple sound backgounds**.** *The Journal of the Acoustical Society of America,* 45(3), 694-703.

Clopper, C. G., & Bradlow, A. R. (2008). Perception of dialect varation in noise: Intelligibility and classification. *Language and Speech*, 51, 175-198.

Cook, Victoria (2009). Speech-intelligibility in cross-dialectal multi-talker babble. BA Seniors honors Thesis, The Ohio State University.

Kidd, G., Mason, C.R., Richards, V.M., . Gallun & Durlach, N.I. (2007). Informational masking. In William A Yost, Arthur N Popper, &, G.J., Richard R. Fay (Eds.),
Auditory perception of sound sources (pp. 143-144-189). New York: Springer.

Hoen, M., Meunier, F., Grataloup, C., Pellegrino, F., Grimault, N., Perrin, F., Perrot, X., Collet L. (2007). Phonetic and lexical interferences in informational masking during speech-in-speech comprehension. *Speech Communication, 49*(12), 905-916.

Jones, C., Berry, L., Stevens, C. 2007. “Synthesized speech intelligibility and persuasion: Speech rate and non-native listeners,” *Computer Speech and Language*. 641-651.

Liu, S., Zeng, F. G. 2006. “Temporal properties in clear speech perception”. *The Journal of the Acoustical Society of America*. 424-432.

Simpson, S. A., and Cooke, M. 2005. “Consonant identification in N-talker babble is a nonmonotonic function of N,” *The Journal of the Acoustical Society of America*. 118, 2775-2778.

Studebaker, G.A. (1985) A "Rationalized" Arcsine Transform. *Journal of Speech and Hearing Research, 28,* 455-462.

Van Engen, K. and Bradlow, A. R. (2007). Sentence recognition in native- and foreign-language multi-talker background noise. *Journal of the Acoustical Society of America*, 121(1), 519-526.

APPENDIX

The following sentences were recorded by 2 male speakers from Western North Carolina (a relatively fast speaker (NC1) and a relatively slow speaker (NC2)). Each set was reordered randomly before being presented to the participants. These sentences were taken from lists 3, 4, 6, 7, 9, 13, and 18 of the Revised Bamford-Kowal-Bench Standard Sentence Test. Scoring of the perception test was based on the number of keywords the listener correctly identified. These keywords are underlined.

Practice

The child grabbed the toy.
They watched the movie.
The tomatoes grew in his garden.
It is time to go home.

Set 1

1. The book tells a story.
2. The young boy left home.
3. They are climbing the tree.
4. She stood near her window.
5. The table has three legs.
6. A letter fell on the floor.
7. The five men are working.
8. He listened to his father.
9. The shoes were very dirty.
10. They went on a vacation.
11. The baby broke his cup.
12. The lady packed her bag.
13. The dinner plate is hot.
14. The train is moving fast.
15. The child drank some milk.
16. The car hit a wall.

Set 2

1. The bakery is open.
2. They like orange marmalade.
3. Mother shut the window.
4. He is skating with his friend.
5. The apple pie was good.
6. Rain falls from the clouds.
7. She talked to her doll.
8. They painted the wall.
9. The towel dropped on the floor.
10. The dog is eating some meat.
11. A boy broke the fence.
12. The yellow pears tasted good.
13. The police helped the driver.
14. The snow is on the roof.
15. The lady washed the shirt.
16. The cup is hanging on a hook.

Set 3

1. The paint dripped on the ground.
2. Mother stirred her tea.
3. They laughed at his story.
4. Men wear long pants.
5. The small boy was asleep.
6. The lady went to the store.
7. The sun melted the snow.
8. The father is coming home.
9. She had her spending money.
10. The truck drove up the road.

11. He is bringing his raincoat.
12. A sharp knife is dangerous.
13. They took some food.
14. The smart girls are reading.
15. The broom stood in the corner.
16. The woman cleaned her house.

Set 4

1. The children dropped the bag.
2. The dog came back.
3. The floor looked clean.
4. She found her purse.
5. The fruit is on the ground.
6. Mother got a saucepan.
7. They washed in cold water.
8. The young people are dancing.
9. The bus left early.
10. They had two empty bottles.
11. The ball is bouncing very high.
12. Father forgot the bread.
13. The girl has a picture book.
14. The orange was very sweet.
15. He is holding his nose.
16. The new road is on the map. 3

Set 5

1. The fruit came in a box.
2. The husband brought some flowers.
3. They are playing in the park.

4. She argued with her sister.
5. A man told the police.
6. Potatoes grow in the ground.
7. He is cleaning his car.
8. The mouse found the cheese.
9. They waited for one hour.
10. The big dog was dangerous.
11. The strawberry jam was sweet.
12. The plant is hanging above the door.
13. The children are all eating.
14. The boy has black hair.
15. The mother heard the baby.
16. The truck climbed the hill.

Set 6

1. The glass bowl broke.
2. The dog played with a stick.
3. The teapot is very hot.
4. The farmer keeps a bull.
5. They say some silly things.
6. The lady wore a coat.
7. The children are walking home.
8. He needed his vacation.
9. Milk comes in a carton.
10. The man cleaned his shoes.
11. They ate the lemon pie.
12. The boy is running away.
13. Father looked at the book.

14. She drinks from her cup.

15. The room is getting cold.

16. A girl kicked the table.